

ROBUST THERMAL FACE RECOGNITION USING REGION CLASSIFIERS

AYAN SEAL^{*,†}, DEBOTOSH BHATTACHARJEE^{*,§},
MITA NASIPURI^{*,¶} and CONSUELO GONZALO-MARTIN^{†,||}

**Department of Computer Science and Engineering
Jadavpur University, India*

*†Center for Biomedical Technology
Universidad Politecnica de Madrid, Spain*

‡ayanseal30@gmail.com

§debotoshb@hotmail.com

¶mitanasipuri@yahoo.com

||consuelo.gonzalo@upm.es

This paper presents a robust approach for recognition of thermal face images based on decision level fusion of 34 different region classifiers. The region classifiers concentrate on local variations. They use singular value decomposition (SVD) for feature extraction. Fusion of decisions of the region classifier is done by using majority voting technique. The algorithm is tolerant against false exclusion of thermal information produced by the presence of inconsistent distribution of temperature statistics which generally make the identification process difficult. The algorithm is extensively evaluated on UGC-JU thermal face database, and Terravic facial infrared database and the recognition performance are found to be 95.83% and 100%, respectively. A comparative study has also been made with the existing works in the literature.

1. Introduction

Automatic face recognition has been comprehensively studied for more than four decades since face recognition of individuals has many applications, particularly in human-machine interaction and security domain. The nonintrusive nature makes the face recognition superior to other biometric-based recognition system is such as fingerprints, iris, hand geometry, etc. Most of the available face recognition systems are based on visible face images. For more information regarding visible face

recognition methods, readers are encouraged to see Ref. 4. These are very good constrained environment. The primary problem in recognizing people in unconstrained environments is the high variability of the visual aspect of the face presented by different sources. The appearance of the face image is changed due to the changes in illumination, occlusions, the head pose and orientation of the subject, the facial expression, and aging. This high changeability makes the face recognition one of the toughest problems in the fields of pattern recognition, computer vision and biometrics. Some of the researchers have been recommended illumination invariant face recognition algorithm based on visible face images.^{24,35} But, still these methods are not sufficient for unconstrained environments. Face recognition using infrared images has made much progress during the last decade with the advent of CCD technology.⁵⁰ Face recognition system based on visual images has an edge over other biometric traits due to easy availability of low cost visible band optical camera, but visual images-based face recognition alone has not been able to achieve the required accuracy because of its sensitivity to illumination, pose, and expressions. On the other hand, thermal infrared face recognition systems do not depend on illumination condition. This system can capture the radiated energy from the face only. As the heat pattern is emitted from the face surface itself, without any source of external radiation, these systems can capture images despite any external lighting and even in the dark. In this system, the tasks of face detection, localization, and segmentation are relatively easier and more reliable than those in visual images.^{25,32} A brief review of the significant works, as found in the literature, on thermal face recognition, as well as infrared face recognition during the last decade, is presented here. In 1992, first thermal face recognition work was conducted in Mikos Corporation.²⁸ Mikos Corporation first demonstrated that the facial thermograms are different for every individual. Wilder *et al.*⁴⁹ have presented one of the first works on thermal face recognition. They have compared the relative performances of three face recognition algorithms for visible and thermal IR images of 101 subjects and concluded that the recognition results for one modality were not significantly better than those for the other. They did not consider a significant amount of illumination changes for visible images and temperature changes for IR images. However, the illumination problem could be solved to a great extent using Gabor transform.⁴⁵ Illumination does not affect the thermal images since thermal images are illumination independent. Cutler¹⁵ were one of the first researchers who have used the eigenface³⁴ method in the infrared face recognition and applied to a database of 288 hand-aligned low-resolution (160×120) images of 24 subjects taken from three viewpoints (frontal, 45° and profile). Recognition rates reported were 96% for frontal views, 96% for 45° views, and 100% for profile views. Another face recognition system based on thermal face was introduced by Yoshitomi *et al.*⁵² in 1997. They focused on the lighting problem in visual face recognition system and recommended infrared face recognition as a solution. Socolinsky *et al.*⁴⁴ presented an illumination invariant face recognition system in Long-Wave Infrared imagery

spectrum (LWIR). They have tested the eigenfaces³⁴ and the ARENA⁴⁸ algorithms on a database of visible and infrared images of 91 distinct subjects taken under different illumination conditions, with varying facial expressions, and with or without glasses. Selinger and Socolinsky⁵² evaluating the performance of the recognition system using the same database of 91 subjects and tested the performance on four different face recognition algorithms (PCA, LDA, LFA and ICA). Their results showed that LWIR imagery of human faces is better to visible imagery in recognition. In Ref. 44, the researchers have used visible imagery and thermal infrared imagery for face recognition in operational scenarios of both indoor and outdoor. They have studied the performance of the face recognition system across multiple sessions. The visible imagery is affected by changes in outdoors illumination; thermal imagery is affected by some challenging factors such as physical activity of subjects, weather conditions for both indoors and outdoors, etc. Their experimental results proved that, under controlled lighting conditions, visible imagery has a better recognition rate for indoors. For all of the algorithms that they used, outdoor recognition rate was worse for both visible and thermal images, with a sharper degradation for visible one. They have concluded that face recognition performance with thermal infrared imagery is steady over many sessions. IR imagery represents a feasible substitute to visible imaging in the search for a robust and practical face recognition system. Trujillo *et al.*⁴² proposed an unsupervised local and global feature extraction paradigm to the problem of facial expression using thermal images. Chen *et al.*⁴⁰ developed a face recognition method with Principle Component Analysis (PCA). They used PCA to study the comparisons and combination of infrared and visible images to the effects of lighting, facial expression change and the time difference between gallery and probe images. They have paid attention to a very crucial factor in thermal infrared face recognition which is the time difference between training and test images in the galleries. First they did not consider elapse time for testing the system. They have stated that there is no regular difference between the recognition rate of visible and thermal IR Infrared. Later, they have concentrated on elapse time. They have concluded that visible image is better than thermal IR in time-lapsed recognition. But there is an important clue in their research: The sets of mismatched probes of the two classifiers for the two models did not overlap necessarily. This clue suggested that these two models may have complementary information about the probe. So, the researchers combined the results of the thermal IR and visible image-based face recognition systems using some combination rules using three different strategies. Their experimental results indicated that when there was significant time elapse between the gallery and probe images, the visible face recognition outperforms the thermal IR one. Their experimental results also showed that the combination of the two models outperforms both the single models. Friedrich *et al.*⁴³ have also proposed an eigenface-based face recognition using thermal IR images. In their work, they found that the thermal IR images are less affected by changes of pose or facial

expression. They found that the thermal IR face recognition is quite better than a visible one when the probe images include highly varying pose and facial expressions. Prokoski *et al.*⁴⁷ expected that it is possible to extract the vascular network i.e. network of blood vessels from thermal facial images as a feature for face recognition. Some of the researchers^{10–12,16,33} used this clue and presented an approach for face recognition based on this physiological information extracted from the thermal face images. First, they segmented the human face from the background using a different method which is mentioned in their papers. Then, they have used to extract blood vessels available on the segmented face. The extracted vascular network is unique for each person. It has been found that there exists an analogy between thermal imprints of human faces and fingerprints of human beings. The thermal imprints of the blood vessels may be treated as the ridges in the fingerprints and fingerprints recognition techniques may be applied on thermal imprints of the human faces for their recognition. The reliability of the fingerprint recognition system has been significantly enhanced through the technique of minutiae extraction from ridges. Most common type of minutiae is: when a ridge either comes to an end, which is called a ridge-termination or when it splits into two ridges, which is called a ridge-bifurcation. So, the number and locations of the minutiae vary from face to face in any particular person. When a set of face images is obtained from an individual person, the number of minutiae is recorded for each face. The precise locations of the minutiae are also recorded, in the form of numerical coordinates, for each face and the feature sets thus obtained has been stored in a separate database. A computer can rapidly compare this feature sets with that of anyone else in the world whose face image has been scanned. This work had a drawback: The matching method could not overcome the nonlinearities in the deformation of the vascular network, due to variations in facial pose and expressions and the experimental results showed high false acceptance rates due to weaknesses of the proposed method in the feature extraction and matching algorithms.

Blood perfusion data can be used for face recognition since blood vessels transporting warm blood are reflected in the thermal patterns. Wu *et al.*³⁶ used this suggestion to get better performance of the IR face recognition system under different surrounding temperatures. First they have extracted the blood perfusion data from thermal patterns, since thermal images are dependent on surrounding temperature and blood perfusion data is more steady features. So, the extracted blood perfusion data (also called physiological data) is used for face recognition. After converting the thermal data into the blood perfusion data, the features are extracted by the PCA methods and RBF neural network is used for recognition. The experimental results have proved that the performance of the face recognition system using blood perfusion data is significantly better than the temperature data. Bhowmik *et al.*³⁷ used Log-polar transform to achieve rotation and to scale invariant images. After that, PCA is used on these transformed images for feature extraction and

finally, classification is done using multilayer perception. When only one-sided semi-profile thermal images of an individual are available, a mosaicking technique can be applied to build an apparent 2D front profile view of that person.³⁹ In Ref. 13, authors have used Log-Gabor wavelets for IR face recognition. In the first step of their used method, Log-Gabor has been used to form feature vector from IR face image. In the second step, PCA is used to reduce the dimensions of the feature vectors that are obtained from the first step. Then Independent Component Analysis (ICA) has been used to form a new feature vector, which fed into Top-match classifier for identifying class labels of the testing IR face images. They have used LWIR face database collected by Equinox Corporation. The available semi-profile image is converted into a mirror image which is similar to the opposite half semi-profile thermal image of that person. Human face is symmetric. So, simple concatenation of these two images (one is the original side view image, and another is the mirror or opposite side view image) produces an apparent 2D front face thermal mosaicked image of that person. This mosaic image can be used in any simple thermal face recognition system.

From the above paragraph, it is clear that most of the researchers have used PCA for extracting features from the face in thermal face recognition. PCA is known as reconstructive approach, which is considered to be robust for contaminated pixels.⁵¹ One of the main deficiencies of the PCA-based classifier is its limited capability to handle local variations in the faces, caused by expressions, motion, deformation,⁸ and false exclusion of thermal information produced by the presence of inconsistent distribution of temperature statistics (missing data), etc.

In this work, decision level fusion²⁷ of many region classifiers is used for the identification of the class label and singular value decomposition (SVD) is used for feature extraction and classification purpose. Region classifier can handle local variations by dividing the face into a number of regions, performing recognition on each of these separate regions and then fusing the results. Due to occlusion or expression change, some of the region classifiers may fail to identify the person. But, the aggregation of decisions of other region classifiers may recognize the person with high confidence. Thus this method is robust against local variations in the faces, caused by occlusions, expression changes, motion, deformation, and false exclusion of thermal information produced by the presence of inconsistent distribution of temperature statistics. Again most of the researchers have implemented their algorithms on few numbers of subjects. Here, we have developed our algorithm on UGC-JU thermal face database³⁰ which contains 84 subjects and which is freely available on request for research purpose. However, the pathological (like fever, headache, inflammation, etc.) and psychological (like nervousness, blush, etc.) conditions of the person, which may affect his blood perfusion data, have not been considered in the present work.

The rest of this paper is organized as follows. The proposed system is introduced in Sec. 2. Experimental results and discussion are presented in Sec. 3. Section 4 describes the comparison with the existing works and Sec. 5 draws the conclusion.

2. Proposed System

A schematic diagram of the proposed Robust Thermal Face Recognition System (RTFRS) is shown in Fig. 1. The system consists of four main modules, namely image acquisition, image preprocessing, region classifier and SVD. Each of these modules has been discussed, in detail, in subsequent subsections.

Face image acquisition: In the present work, thermal and visible face images are acquired simultaneously with variable expressions, poses and with/without glasses. Till now 84 individuals have volunteered for this photo shoots and each of 39 different templates of RGB color images with different expressions (happy, angry, sad, disgusted, neutral, fearful and surprised) and pose changes about x-, y- and z-axis are taken. Resolution of each image is 320×240 , and the images are saved in JPEG format. FLIR 7 camera is used to capture this database. Our FLIR 7 camera can capture thermal infrared spectrum (wavelength of $8\text{--}14\ \mu\text{m}$). A typical thermal face image is shown in Fig. 2(a)). This thermal face image depicts interesting thermal information of a facial model.

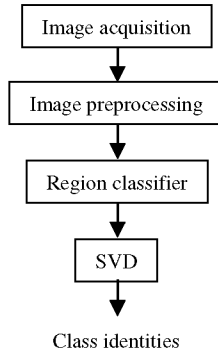


Fig. 1. Schematic block diagram of the proposed system.

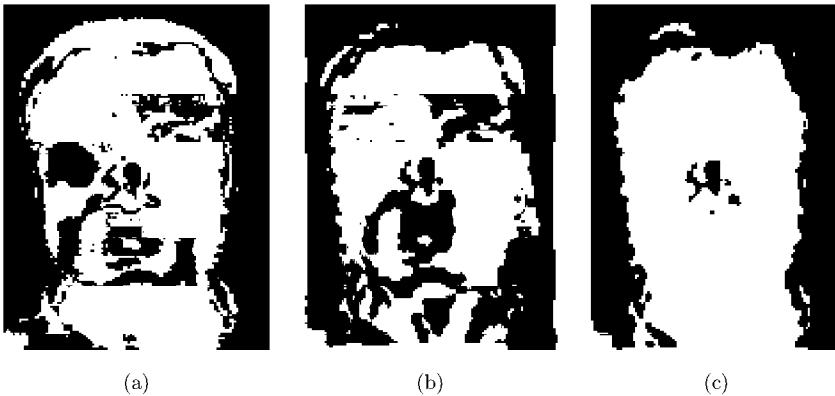


Fig. 2. Thermal face image and its various preprocessing stages. (a) Athermal face image, (b) corresponding grayscale image and (c) binary image.

2.1. Image preprocessing

The stage involves binarization of the acquired thermal face image, extraction of the largest component as the face region, and cropping of the face region in elliptic shape. The image processing techniques which have been used here are discussed in detail as follows.

2.2. Binarization

Each of the captured 24-bits color images has been converted into its 8-bit grayscale image.⁴¹ The grayscale image of the sample image of Fig. 2(a) is shown in Fig. 2(b). The grayscale images are then converted into corresponding binary images by setting pixels in the grayscale image with luminance greater than the mean intensity with the value 1 (white) and all other pixels with the value 0 (black). In a binary image, black pixels mean background and white pixels mean the face region. The binary image of the grayscale image given in Fig. 2(b) is shown in Fig. 2(c).

Extraction of the largest component⁴¹: The foreground of a binary image may contain more than one object. Let us consider the image, in Fig. 2(c), it has six foreground objects or components. The largest one represents the face region. The others are at the left bottom corner and small dot on the top. Then largest component has been extracted from a binary image using “Connected Component Labeling” algorithm.⁹ This algorithm is based either on “4-connected” neighbors or “8-connected” neighbor’s method.^{7,17} In this study “8-connected” neighbor’s method is used. It is a binary image. Here, white means face skin region and black means background. Using the selected largest component as template face region from the grayscale image is then extracted. Figure 3(b) shows only the gray level face region (excluding background) of Fig. 3(a).

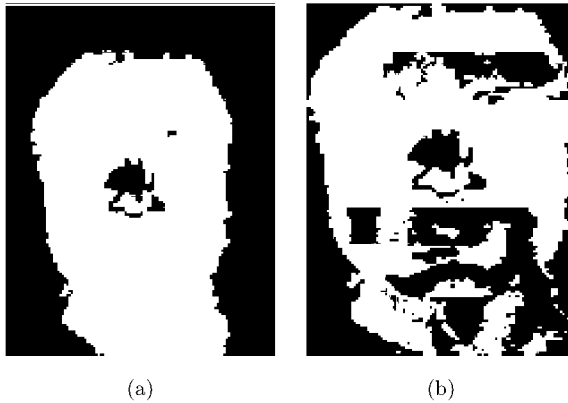


Fig. 3. (a) A largest component as a face skin region. (b) A largest component as a face skin region (grayscale image).

2.3. Cropping of the face region in elliptical shape

Normally human face is of an elliptical shape. For fitting an ellipse on the identified face region, the centroid^{29,41} of the face region needs to be evaluated. The coordinates of the centroid are evaluated from the binarized face image using the following equations.

$$X = \frac{\sum f(x, y)x}{\sum f(x, y)}, \quad Y = \frac{\sum f(x, y)y}{\sum f(x, y)},$$

where x, y are the co-ordinates of the binary image and $f(x, y)$ is the intensity value at x, y co-ordinates. Using this information, face region has been cropped in an elliptical shape using “Bresenham ellipse drawing”¹⁸ algorithm. Distance between the centroid and the right ear is called the minor axis of the ellipse and distance between the centroid and the forehead is called the major axis of the ellipse. The cropped elliptical face image is shown in Fig. 4.

2.4. Region classifier

After cropping of the face region in elliptical shape, it has been seen that some thermal information within the facial area (especially nose and eye region) is excluded during binarization, shown in Fig. 4. False exclusion of thermal information produced by the presence of inconsistent distribution of temperature statistics makes the process of identifying persons using their thermal facial images difficult. So, in the classification stage, we propose to consider the thermal face surface as a combination of several local regions. If the facial area is partially excluded due to some reasons, information regarding the excluded regions will not be available. Therefore instead of globally evaluating a surface, it will be beneficial to analyze the images by separate local classifiers and then to fuse the regional classification results. Local analysis of



Fig. 4. Face region in elliptic shape.

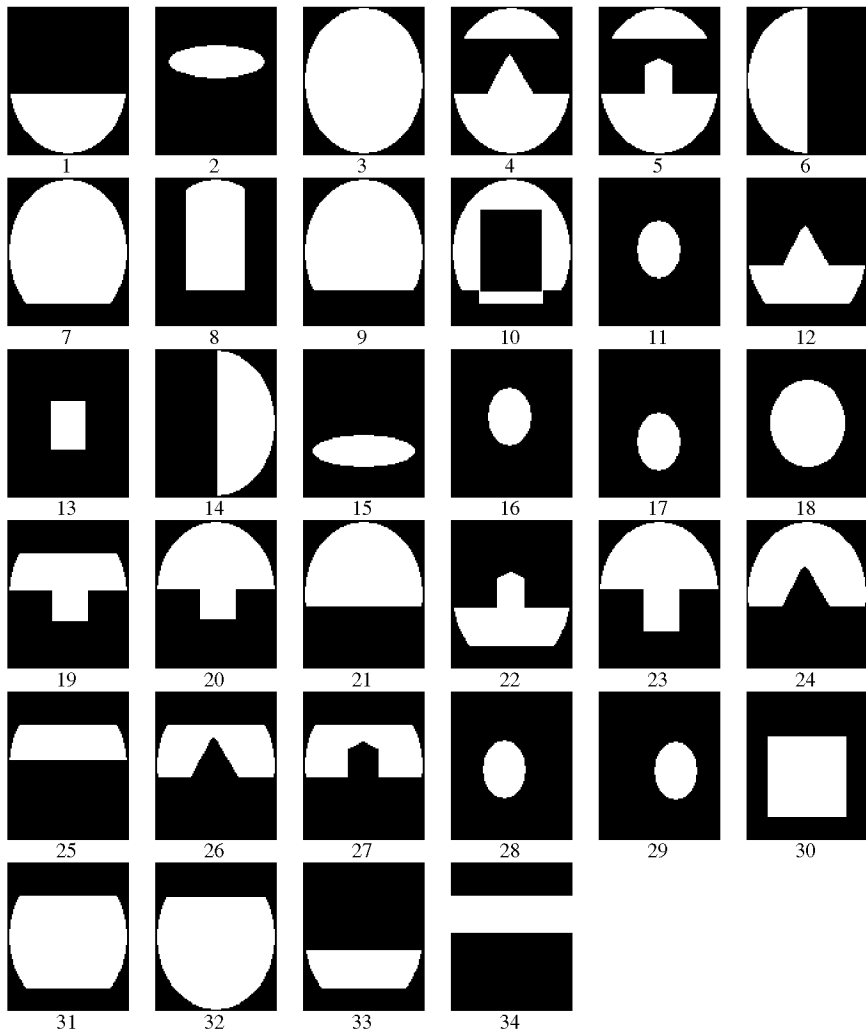


Fig. 5. Local region templates.

the facial surface was proposed in Ref. 1 to deal with facial occlusions on 3D images, which was utilized in Ref. 2 for facial expression variations. In this work, we utilized 34 overlapping local region templates of the face, as shown in Fig. 5, where the white area represents the selected local face region and the black area is the excluded face region. The regions were chosen in such a way that, for different types of local variation, they would allow stable features for comparison. Examples of such regions are those that leave out the upper or the lower part of the face because of the difference in hair, caps, etc. or difference in expression of the mouth. Other examples are leaving out areas covered by glasses and the left or right side of the face, which are less visible for large rotations around the vertical y -axis.

All of these local regions are mapped on the elliptical face, and the features are extracted from the white region only. All feature vectors can be represented as a $1 \times n$ -dimensional vector. So, the size of the feature vector is less compared to that extracted from the whole face image. Still the size of the feature vector is not manageable. Since, any attempt at recognition in such high-dimensional space is vulnerable to a variety of issues often referred to as the curse of dimensionality. Therefore, at the feature extraction stage, images are transformed to low-dimensional vectors in the face space. The main objective is to find such a basis function for this transformation which could distinguishably represent faces in the face space. A number of approaches have been reported in the face recognition literature, such as PCA,^{23,34} LDA,^{5,14} ICA,^{3,21} SVD⁶ and Wavelet transforms.^{29,31,46}

PCA is known as reconstructive approach. Reconstructive approach is robust for contaminated pixels⁵¹ which could serve our purpose. In this work, instead of solving through covariance matrix, eigenvectors and the corresponding eigenvalue are found out through SVD because the covariance matrix-based method is computationally as well as memory requirement wise happens to be very costly.

Singular Value Decomposition: SVD is considered as an important topic in linear algebra by many renowned mathematicians. It is the widely used technique to decompose a matrix into several component matrices, exposing many useful and interesting properties of the original matrix. The decomposition of a matrix is often called a factorization. Let us say A be a rectangular matrix with m rows and n columns, with rank r and $r \leq n \leq m$. Then A can be broken down into the product of three matrices — an orthogonal matrix U , a diagonal matrix S , and the transpose of another orthogonal matrix V :

$$A_{mn} = U_{mn} S_{nn} V_{nn}^T,$$

where $U^T U = I$, $V^T V = I$; the matrix $U = [u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_m]$ are orthonormal eigenvectors of $A^T A$, and the column vector u_i , for $i = 1, 2, \dots, m$, form an orthonormal set.

$$u_i^T u_j = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Similarly, the matrix $V = [v_1, v_2, \dots, v_r, v_{r+1}, \dots, v_n]$ contains orthonormal eigenvectors of $A^T A$, the column vector v_i , for $i = 1, 2, \dots, n$, form an orthonormal set:

$$v_i^T v_j = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

and S is a diagonal matrix with singular values $\sigma_1, \sigma_2, \dots, \sigma_n$ of A on the diagonal which are the square roots of eigenvalues obtained from U or V in descending order.

So, SVD allows us to factoring a digital image in three matrices. The use of S_{nn} , i.e. singular values allows us to represent the image with a smaller set of values,

which can preserve useful features of the original image, but use less storage space in the memory. PCA can be derived from the SVD.²⁶ The covariance matrix can be written as

$$C = \frac{1}{m} AA^T = \frac{1}{m} US^2V^T.$$

From the above equation, it is understood that SVD order the singular values in decreasing order and the first m columns corresponds to the sorted nonzero eigenvalues of C as $m > n$. So, PCA transformed data can be written as

$$X = \bar{U}^T A = \bar{U}^T U S V^T,$$

where \bar{U} is the matrix of eigenvectors sorted according to eigenvalues and $\bar{U}^T U$ is a simple $m \times n$ matrix which is one on the diagonal and zero everywhere else. So, the transformed PCA data can be described in terms of the SVD decomposition of A .

SVD for face recognition: Here, a single face image in the training phase is known as *baseface*, and all such *basefaces* are grouped together to create a *face space* which are known faces. When, a new face image appears for identifying its class label by projecting it onto the face space, and then matching this face by comparing its coordinates (position) in face space with the coordinates (positions) of known faces i.e. “face space”. Let I be a *baseface* of size m rows and n columns. A row vector f_i of size $1 \times M$ ($= m \times n$) has been created by appending each row one after another column wise. Each such *baseface* is appended row wise in order to get the training set S with N number of face images of known individuals to be represented in the form of an $N \times M$ matrix:

$$S = [f_1, f_2, \dots, f_N]^T.$$

The mean image f_{mean} of training set S is found by $f_{\text{mean}} = \frac{1}{N} \sum_{i=1}^N f_i$ which is shown in Fig. 6.

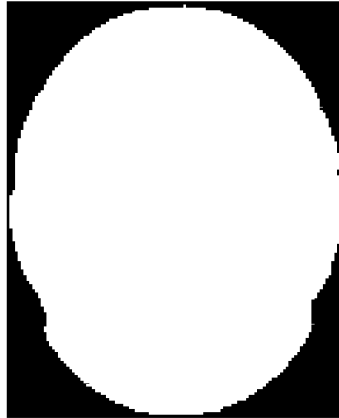


Fig. 6. Mean face of the face images in the training set.

Each of the training images must be centered. This centering is done by subtracting the mean image from each of the training images as shown below:

$$x_i = f_i - f_{\text{mean}}, \quad i = 1, 2, \dots, N.$$

This gives another $M \times N$ matrix X :

$$X = [x_1, x_2, \dots, x_N].$$

Then SVD is applied on the matrix X as described below:

$$X_{MN} = U_{MM} S_{MN} V_{NN}^T.$$

Depending on the nonzero singular values $\{\sigma_1, \sigma_2, \dots, \sigma_r\}$, $\{u_1, u_2, \dots, u_r\}$ form an orthonormal basis for $R(X)$, the range (row) subspace of matrix X where $R(X)$ is called a *face subspace* in the ‘image space’ of $m \times n$ pixels, and each u_i , $i = 1, 2, \dots, r$, can be called a *baseface*.

The input image $f_{m \times n}$ in the *facespace* is represented by the contribution of each *baseface* which are treated as basis as follows:

$$x = [u_1, u_2, \dots, u_r]^T \times (f - f_{\text{mean}}).$$

The vector x will be used in recognition to find which predefined face classes are the best matches of the face. The simplest method for determining which face class provides the best description of an input class x_i is through the use of Euclidean distance and trying to find which face class minimizes it.

$$d_i = |x - x_i| = \sqrt{(x - x_i)^T (x - x_i)},$$

where x_i is the feature vector of f_i , which is the scalar projection of $(f_i - f_{\text{mean}})$ onto the “*baseface*”:

$$x_i = [u_1, u_2, \dots, u_r] \times (f_i - f_{\text{mean}}).$$

A face f is classified as f_i when d_i is found to be minimum.

So, the recognition process starts from projecting a region template onto the training images and extracting the pixel information from the mapped region (white area) of the training images, then storing these pixels’ information as a row vector which is treated as a feature vector. Size of this feature vector depends on the white area of the region template. It has been seen that the size of the feature vector is very large. Combining these feature vectors as columns a matrix is formed. Application of SVD transforms this matrix into a product USV^T , which is refactoring of a digital image in three matrices. This refactoring can conserve useful features of the original image and represents the image with a smaller set of values instead of using less storage space in the memory. The singular values $\sigma_1, \sigma_2, \dots, \sigma_n$ are unique; however the matrices U and V are not unique, so each image has its own singular values. After extracting SVD for the train and test sets, by using Euclidean distance, minimum distance will be found.

The recognition results coming from different region classifiers are independent and may vary widely. Some of them may be very low. But fusion of the results of the region classifiers into a single score or decision might be better than the individual decisions.

In our work, decision level fusion being adopted is based on majority voting technique since majority voting technique will be more appropriate with the idea of using multiple region classifiers each representing more or less stable regions in the face area. Some of the region classifiers may give incorrect decision, but still many others will give the correct decision. So, majority voting technique takes decision when the majority of the classifiers agree.

3. Experimental Results and Discussion

Experiments have been performed on UGC-JU thermal face database,³⁸ and one benchmark database named as Terravic facial infrared database. The details of UGC-JU face database have been mentioned earlier. Twelve frontal images are taken for each person for our experiments from the two datasets i.e. our UGC-JU thermal face dataset and Terravic facial infrared dataset separately. All the images of size 112×92 have been made. Then these thermal face images are divided into two sets, namely training set and test set. Here two types of experiments have been performed. In the first set of experiments, the entire face image is used for classification. The pixel information from the cropped elliptic thermal face images of each (person) in the training set is fed to the SVD module, which defines the *basefaces* and *facespace* for the training set. For testing, each face image in the test set is projected onto the face-subspace of the each training images, and its distance from the nearest base-face of the training set is computed. This is repeated for all the test images and the test image is assigned with the class label for which that distance measure is found minimum. A number of experiments have been performed, which are discussed in detail below:

3.1. 2-fold cross-validation

In the 2-fold cross-validation method, total image set is divided into two disjoint sets U and V , respectively, and the U set is used for training and V set is used for testing purpose and the vice versa. Then the average percentage of recognition rate has been calculated on UGC-JU thermal face database, and the obtained result is shown in Table 1. The average recognition rate is found to be 100% for Terravic facial infrared database.

Table 1. Performance rate for simple approach using 2-fold cross-validation.

Training Set	Testing Set	Recognition Rate (%)
U	V	91.67
V	U	88.89
Average recognition rate (%)		90.28

3.2. 3-fold cross-validation

In the 3-fold cross-validation method, total images are divided into three disjoint sets U , V and X , respectively. In the first part of this experimentation (Exp 1), one of these three sets is used for training purpose and rest two sets are used for testing purpose. The recognition rate on the test sets is calculated and shown in Table 2. The average percentage of recognition rate is found to be 100 for Terravic facial infrared database.

In the second part of experiment (Exp 2), two sets are used for training purpose and rest one set is used for testing purpose. The obtained recognition rates on the test set are calculated and reported in Table 3. The average recognition rate is found to be 100% for Terravic facial infrared database.

In the second set of experiments, region classifiers are used. Here, total images are also tested using both 2-fold and 3-fold cross-validation. For feature extraction, each region template is projected onto the training images, and pixel information is extracted from the mapped area (white area) from the training images. The pixel information is fed to the SVD module, and *facesubspace* and *baseface* are defined. So, this process is repeated for all the region templates (34 different regions). The performance of individual classifiers and the fused result for 2-fold cross-validation are shown in Fig. 7. 'X'-axis of the following graphs represents the region classifier number and 'Y'-axis represents the recognition rate in percentage. Here, the total classifier numbers are 34. So, in 'X'-axis 34 different marks are there. The 35th mark is used to represent the fused and average result. Five different symbols (diamond, square, triangle, multiplication and asterisk) are used to represent the results of Training Set U and Testing Set V , Decision level fusion (f1), Training Set V and Testing Set U , Decision level fusion (f2) and Average recognition rate respectively

Table 2. Performance rate for simple approach using 3-fold cross-validation (Exp 1).

Training Set	Testing Set	Recognition Rate (%)
U	$V + X$	85.42
V	$U + X$	87.50
X	$U + V$	85.42
Average recognition rate (%)		86.11

Table 3. Performance rate for simple approach using 3-fold cross-validation (Exp 2).

Training Set	Testing Set	Recognition Rate (%)
$V + X$	U	87.50
$U + X$	V	87.50
$U + V$	X	91.67
Average recognition rate (%)		88.89

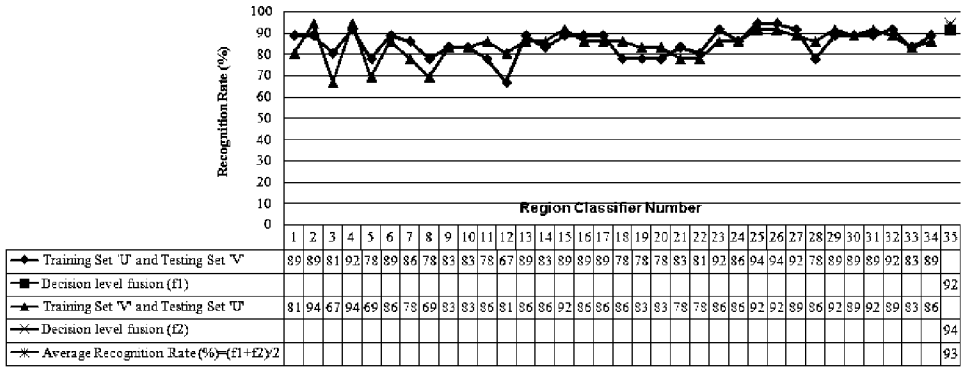


Fig. 7. Region classifier number versus recognition rate (%) for 2-fold cross-validation.

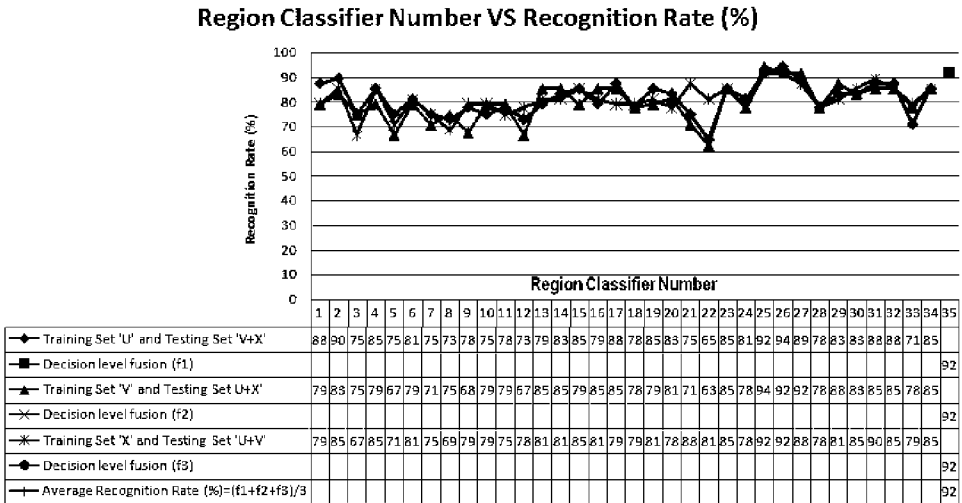


Fig. 8. Region classifier number versus recognition rate (%) for 3-fold cross-validation (Exp 1).

and the average Recognition rate is found to be 93.06%. The average recognition rate is found to be 100% for Terravic facial infrared database.

In the 3-fold cross-validation method, total image set is divided into three disjoint sets U , V and X , respectively. The 3-fold cross-validation method is also divided into two ways. In the first way, one set out of three sets is used for training and rest of the two sets is used for testing purpose, and the result is shown in Fig. 8. Average recognition result is 91.67%. In the second way, two sets are used for training and one set is used for testing purpose. Result is shown in Fig. 9. Average recognition result is 95.83%. The average recognition rate is found to be 100% for Terravic facial infrared database.

From the above tables, we have two little observations. First one is that, the region classifiers give a better result than a simple approach (without a region

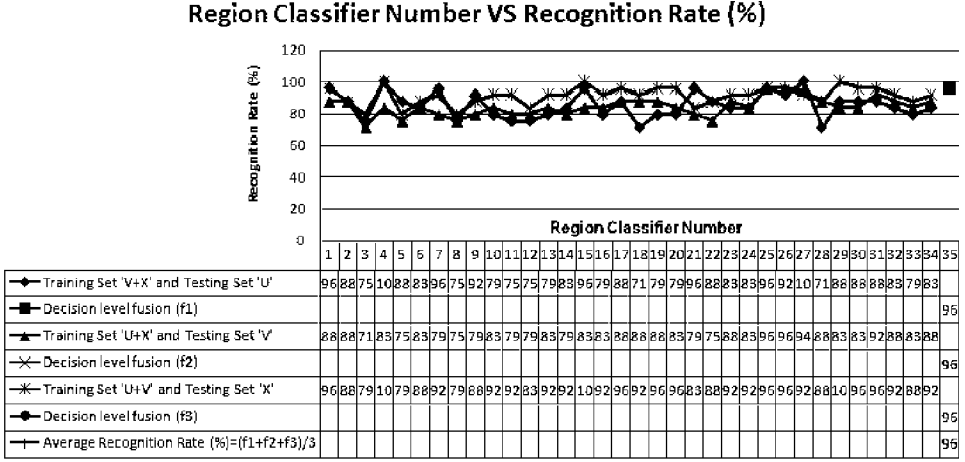


Fig. 9. Region classifier number versus recognition rate (%) for 3-fold cross-validation (Exp 2).

classifier). Human can identify or distinguish a person by seeing his or her eye regions only since it is have a unique characteristic. The recognition results obtained from the Region classifier numbers 25, 26, and 27, which are concentrated on eye regions only, are better than that obtained from the other parts of the face.

4. Comparison with Other Methods

The present method has been compared with previous works of Seal *et al.*^{36,37,39} on thermal face image recognition on the basis of their performances on the UGC-JU face database. In their previous work, three different methods have been used to extract the blood perfusion image, namely bit-plane slicing and medial axis transform, morphological erosion and medial axis transform, 'sobel' edge operators. Then two methods have been used to find the minutiae points from a blood perfusion image of a human face, which are unique for that face. After the extraction of minutiae points from a blood perfusion image, entire face image is divided into equal size square blocks, and the total number of minutiae points from each block is calculated and stored it in a vector to make final feature vector. Therefore, the size of the feature vectors is found to be same as the total number of blocks considered. A three layer feedforward back propagation neural network is used as a classification tool. The maximum recognition rate obtained on these methods is 95.24% which is quite satisfactory. But, these methods have several limitations. First, these minutiae-based techniques may be affected by occlusion, wear glass since these techniques are sensitive to proper localization of the minutiae points. Second, the success rate depends on the size of the block. Lastly, the exact localization of minutiae points are always a difficult task. These limitations have been overcome using 34 different Region-based classifiers, each was focusing on different regions of the face. Here, the recognition performance depends on the aggregation of the decisions of the individual

region classifiers; failures of some of them are supplemented by the success of others. The present method has also been compared with one work, which is based on PCA, Wavelet, and SVM.¹⁹ In this work, the authors have used PCA and Haar level-4 wavelet for extraction of features from the face images and fed these features into SVM with three types of kernels namely RBF kernel, Linear polynomial kernel and Quadratic polynomial kernel and Nearest Distance Classifier separately for classification purpose. PCA and Haar level-4 wavelet are used in this work to extract the feature from the UGC-JU thermal face images for comparisons. The size of the low-frequency coefficient array is 7×6 after level-4 decomposition. In this work, only LL4 coefficients are used. The LL4 coefficients of a face image are stored in the form of a row vector, and the size of the row vector is 1×42 ($7 \times 6 = 42$). Here the face images are divided into two groups; one is used for training and the other for testing. All experiments are performed in MATLAB 2012 platform except SVM classification. For SVM, Weka 3.6.10 version is used. The obtained average recognition rate after 2-fold cross-validation is shown in Table 4.

In Ref. 19, authors have considered different pose changes for their work, but in the present implementation only frontal face images with different facial expressions are considered. The present work has also been compared with one of the existing work of Chen and Jing.¹³ The used algorithm consists of four steps. In the very first step, Log-Gabor wavelets are used at four different scales and six directions to form 24 images of the same size of the input face image. Some of the parameters need to be set for executing Log-Gabor wavelet transform. These parameters are wavelength of smallest scale filter, scaling factor between successive filters, ratio of the standard deviations of the radial Gaussian function to the filter center frequency, and ratio of angular interval between filter orientations and the standard deviation of the angular Gaussian function. In this experiment, the value of these parameters are set to be 3, 2, 0.65, and 1.5, respectively. Then all 24 transformed images are concatenated column wise to form feature vector. Then PCA has been used to reduce the size of the feature vector. ICA has been considered to get a new feature vector from the reduced feature vector. Top-match classifier is used for identifying the class label of the test

Table 4. Comparison of recognition performances of different methods on UGC-JU thermal face image database.

Methods	Recognition Rate (%)
Haar Wavelet–SVM (RBF Kernel) ¹⁹	91.67
Haar Wavelet–SVM (Linear polynomial kernel) ¹⁹	87.50
Haar Wavelet–SVM (Quadratic polynomial kernel) ¹⁹	88.89
Haar Wavelet–Nearest Distance ¹⁹	79.00
PCA–SVM (RBF Kernel) ¹⁹	85.42
PCA–SVM (Linear polynomial kernel) ¹⁹	88.42
PCA–SVM (Quadratic polynomial kernel) ¹⁹	87.50
PCA–Nearest Distance ¹⁹	72.22
Log-Gabor wavelet-based method ¹³	88.89
Proposed method	95.83 (Max)

face image. The same experiment has been done on UGC-JU thermal face database, and the obtained result is shown in Table 4. All the results support the conclusion that the face recognition performances based on region classifiers are better than all other above mentioned methods, and the proposed method also helps to remove the limitations of the above mentioned problems.

5. Conclusion

In this work, Region Classifier is used for the first time in the area of thermal face recognition. We have proposed a thermal face recognition system which is robust against the missing thermal information of the facial surface produced by the presence of inconsistent distribution of temperature statistics. Experiment is done on UGC-JU thermal face database and Terravic facial infrared database. Extensive experiments were carried out to illustrate the efficiency of thermal face recognition using Region classifier and SVD. In terms of face classification, we have shown that dividing the facial surface into local regions and combining their individual views at the classification level significantly improves the recognition rate. Furthermore, using SVD-based classifiers improves identification results: when local results are fused by majority voting technique. The simple architecture of the proposed approach makes it computationally efficient. Besides, no knowledge of geometry or specific feature of the face is required. However, this system is applicable to front views and constant background only. It may fail in unconstrained environments like natural scenes.

Acknowledgments

Authors are thankful to a major project entitled “Design and Development of Facial Thermogram Technology for Biometric Security System,” funded by University Grants Commission (UGC), India and “DST-PURSE Programme” at Department of Computer Science and Engineering, Jadavpur University, India for providing the necessary infrastructure to conduct experiments relating to this work. Ayan Seal is grateful to Department of Science & Technology (DST), India for providing him Junior Research Fellowship-Professional (JRF-Professional) under DST-INSPIRE Fellowship programme [No: IF110591].

References

1. N. Alyuz, B. Gokberk and L. Akarun, 3D face recognition system for expression and occlusion invariance, *2nd IEEE Int. Conf. Biometrics: Theory, Applications and Systems (BTAS)* (IEEE, 2008), pp. 1–7.
2. N. Alyuz, B. Gokberk and L. Akarun, Regional registration for expression resistant 3-D faces recognition, *IEEE Trans. Inf. Forensics Security* **5**(3) (2010) 425–440.
3. M. Bartlett, H. Lades and T. Sejnowski, Independent component representations for face recognition, in *Proc. SPIE Conf. Human Vision and Electronic Imaging III* (1998), pp. 528–539.

4. M. P. Beham and S. M. M. Roomi, A review of face recognition methods, *Int. J. Pattern Recogn. Artif. Intell.* **27**(4) (2013) 1356005.
5. P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection, *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7) (1997) 711–720.
6. G. Beylkin, R. Coifman and V. Rokhlin, Fast wavelet transforms and numerical algorithms I, *Commun. Pure App. Math.* **44**(2) (1991), 141–183.
7. D. Bhattacharjee, A. Seal, S. Ganguly, M. Nasipuri and D. K. Basu, A comparative study of human thermal face recognition based on Haar wavelet transform (HWT) and local binary pattern (LBP), *Comput. Intell. Neurosci.* (2012) doi: 10.1155/2012/261089.
8. M. K. Bhowmik, D. Bhattacharjee, M. Nasipuri, D. K. Basu and M. Kundu, Classification of polar-thermal eigenfaces using multi-layer perceptron for human face, *IEEE Region 10 Colloquium and the Third ICIS*, Kharagpur, India, 8–10 December (2008), pp. 1–6.
9. P. Buddharaju, I. Pavlidis and I. Kakadiaris, Face recognition in the thermal infrared spectrum, in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition Workshops (CVPRW '04)* (2004).
10. P. Buddharaju, I. Pavlidis and P. Tsiamyrtzis, Physiology-based face recognition using the vascular network extracted from thermal facial images: A novel approach, in *Proc. IEEE Int. Conf. Advanced Video and Signal Based Surveillance*, Como, Italy, 15–16 September, 2005, pp. 354–359.
11. P. Buddharaju, I. T. Pavlidis, P. Tsiamyrtzis and M. Bazakos, Physiology-based face recognition in the thermal infrared spectrum, *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4) (2007) 613–626.
12. X. Chen, P. J. Flynn and K. W. Bowyer, PCA-based face recognition in infrared imagery: Baseline and comparative studies, in *Proc. IEEE Int. Workshop on Analysis and Modeling of Faces and Gestures (AMFG '03)* (2003), pp. 127–134.
13. X. Chen and Z. Jing, Infrared face recognition based on log-Gabor wavelets, *Int. J. Pattern Recogn. Artif. Intell.* **20**(3) (2006) 351–360.
14. P. Comon, Independent component analysis — A new concept?, *Signal Process.* **36** (1994), 287–314.
15. R. Cutler, Face recognition using infrared images and eigenfaces (1996), Available at <http://citeseer.ist.psu.edu/cutler96face.html>.
16. G. Friedrich and Y. Yeshurum, Seeing people in the dark: Face recognition in infrared images, in *Second BMCV* (2003), pp. 348–359.
17. S. Gong, S. J. McKenna and A. Psarrou, *Dynamic Vision: From Images to Face Recognition* (ImperCooege Press, 2000).
18. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd edn. (Prentice Hall, 2002).
19. E. Gumus, N. Kilic, A. Sertbas and O. N. Ucan, Evaluation of face recognition techniques using PCA, wavelets and SVM, *Exp. Syst. Appl.* **37** (2010) 6404–6408.
20. U. Halici, L. C. Jain and A. Erol, An introduction to fingerprint recognition, in *Intelligent Biometric Techniques in Fingerprint and Face Recognition* eds. L. C. Jain, U. Halici, I. Hayashi, S. B. Lee and S. Tsutsui (CRC Press, Florida, 1999), pp. 3–34.
21. Z. Q. Hong, Algebraic feature extraction of image recognition, *Pattern Recogn.* **24**(3) (1991) 211–219.
22. A. Jain, A. Ross and S. Prabhakar, An introduction to biometric recognition, *IEEE Trans. Circuits Syst. Video Technol.* **14**(1) (2004) 4–20.
23. I. T. Joliffe, *Principal Component Analysis* (Springer, 1986).
24. G. A. Khuwaja and M. S. Laghari, Adaptive classifier integration for invariant face recognition (ACIIFR), *Int. J. Pattern Recogn. Artif. Intell.* **16**(6) (2002) 749–772.

25. S. G. Kong, J. Heo, B. R. Abidi, J. Paik and M. A. Abidi, Recent advances in visual and infrared face recognition—A review, *Comput. Vis. Image Understand.* **97** (2005) 103–135.
26. R. E. Madsen, L. K. Hansen and O. Winther, Singular value decomposition and principal component analysis, February 2004, Available at <http://www.library.cornell.edu/nr/bookcpdf.html>.
27. S. Majumder, G. Majumder and M. Bhowmik, Human face mosaicing using thermal image, in *Journal of the Tripura Mathematical Society*, Special Vol. 13 (Tripura Mathematical Society, 2011), pp. 145–153.
28. C. Manohar, Extraction of superficial vasculature in thermal imaging, Master's thesis, Dept. Electrical Eng., Univ. of Houston, Houston, Texas (2004).
29. B. S. Morse, Lecture 2: Image Processing Review, Neighbors, Connected Components, and Distance (1998–2004), available at <http://homepages.inf.edu.ac.uk/rbf/CVonline/LOCAL.COPIES/MORSE/connectivity.pdf>.
30. I. Naseem, Linear regression for face recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(11) (2010) 2106–2112.
31. OTCBVS WS Series Bench; J. Davis and M. Keck, A two stage approach to person detection in thermal imagery, in *Proc. Workshop on Applications of Computer Vision*, January 2005, pp. 364–369.
32. F. Prokoski, History, current status, and future of infrared identification, in *Proc. IEEE Workshop Computer Vision Beyond Visible Spectrum: Methods and Applications* (2000), pp. 5–14.
33. F. J. Prokoski and R. Riedel, Infrared identification of faces and body parts, in *BIO-METRICS: Personal Identification in Networked Society*, Chap. 9 (Kluwer Academic Publishers, 1998), pp. 1–23.
34. F. J. Prokoski, R. B. Riedel and J. S. Coffin, Identification of individuals by means of facial thermography, in *Proc. IEEE 1992 Int. Carnahan Conf. Security Technology: Crime Countermeasures*, Atlanta, GA, USA, 14–16 October 1992 (IEEE), pp. 120–125.
35. D. Rizo-Rodríguez, H. Méndez-Vazquez and E. García-Reyes, Illumination invariant face recognition in quaternion domain, *Int. J. Pattern Recogn. Artif. Intell.* **27**(3) (2013) 1360004.
36. A. Seal, D. Bhattacharjee, M. Nasipuri and D. K. Basu, Minutiae based thermal face recognition using blood perfusion data, in *Proc. IEEE Int. Conf. Image Information Processing*, Jaypee University of Information Technology, India, 3–5 November, 2011, pp. 1–4.
37. A. Seal, D. Bhattacharjee, M. Nasipuri and D. K. Basu, Minutiae from bit-plane sliced thermal images for human face recognition, in *Proc. Springer Int. Conf. Soft Computing for Problem Solving*, Roorkee, India, 20–22 December, 2011, pp. 113–124.
38. A. Seal, D. Bhattacharjee, M. Nasipuri and D. K. Basu, UGC-JU face database and its benchmarking using linear regression classifier, *Multimedia Tools Appl.* (2013), doi: 10.1007/s11042-013-1754-8.
39. A. Seal, S. Ganguly, D. Bhattacharjee, M. Nasipuri and D. K. Basu, Minutiae based thermal human face recognition using label connected component algorithm, in *Proc. International Conference on Computer, Communication, Control and Information Technology*, India, 25–26 February, 2012, pp. 604–611.
40. A. Selinger and D. A. Socolinsky, Appearance-based facial recognition using visible and thermal imagery: A comparative study, Equinox Corporation, Technical Report 02-01 (2002).
41. Y. Shinohara and N. Otsu, Facial expression recognition using fisher weight maps, *Sixth IEEE Int. Conf. Automatic Face and Gesture Recognition*, Vol. 100 (2004), pp. 499–504.
42. T. Sim, R. Sukthankar, M. Mullin and S. Baluja, Memory-based face recognition for visitor identification, in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition* (2000), pp. 214–220.

43. D. A. Socolinsky and A. Selinger, Thermal face recognition in an operational scenario, in *Proc. CVPR 2004*, Washington, DC (2004), pp. 1012–1019.
44. D. Socolinsky, L. Wolff, J. Neuheisel and C. Eveland, Illumination invariant face recognition using thermal infrared imagery, in *IEEE Computer Society Int. Conf. Computer Vision and Pattern Recognition*, Vol. 1. Kauai, HI, USA (2001), pp. 527–534.
45. L. J. Spreeuwers, Fast and accurate 3D face recognition using registration to an intrinsic coordinate system and fusion of multiple region classifiers, *Int. J. Comput. Vis.* **93**(3) (2011) 389–414.
46. E. J. Stollnitz, T. D. DeRose and D. H. Salesin, Wavelets for computer graphics: A primer, part 1, *IEEE Comput. Graphics Appl.* **15**(3) (1995) 76–84.
47. L. Trujillo, G. Olague, R. Hammoud and B. Hernandez, Automatic feature localization in thermal images for facial expression recognition, in *Proc. CVPR '05 Proc. the 2005 IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR '05)*, Washington, DC, USA (2005), p14.
48. M. Turk and A. Pentland, Eigenfaces for recognition, *J. Cognit. Neurosci.* **3** (1991) 71–86.
49. J. Wilder, P. Phillips, C. Jiang and S. Wiener, Comparison of visible and infra-red imagery for face recognition, in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition (AFGR '96)*, Killington, Vermont, October 1996, pp. 182–187.
50. S. Wu, Z. J. Fang, Z. H. Xie and W. Liang, Blood perfusion models for infrared face recognition, School of information technology, Jiangxi University of Finance and Economics, China, Vienna, Austria (2008), p. 236.
51. S. Wu, W. Song, L. J. Jiang, S. Xian, F. Pan, W. Y. You and S. Ranganath, Infrared face recognition by using blood perfusion data, in *Proc. Int. Conf. Audio Video Based Biometric Person Authentication*, New York, USA, 20–22 July, 2005, pp. 320–328.
52. Y. Yoshitomi, T. Miyaura, S. Tomita and S. Kimura, Face identification using thermal image processing, in *Proc. IEEE Int. Workshop Robot Human. Communication* (1997), pp. 374–379.